

CLASSICAL VS. BAYES RELIABILITY GROWTH IN THEORY AND PRACTICE

Vasily Krivtsov
Ford Motor Co.
20901 Oakwood Blvd.
MD 412
Dearborn, MI 48121-2053

James Wasiloff
Ford Motor Co.
36200 Plymouth Road
MD 32
Livonia, MI 48151

SUMMARY

A comprehensive, succinct review clarifying the merits of classical and Bayesian based Reliability Growth methodologies. Provides practical guidelines for the application of each of the methods discussed.

KEY WORDS

Bayesian, reliability growth, statistical models

INTRODUCTION

As the quest for improved customer satisfaction in today's highly competitive marketplace and its inherent link to high-level, long-term reliability continues to emerge, the importance of a robust method of comparing two designs and/or evaluating the magnitude of reliability improvement between consecutive design iterations is becoming a reality. Product engineering must understand, internalize, and properly apply the "right tool" as required to ensure achieving goals in customer satisfaction.

As the design cycle of a product progresses from the concept to development, testing, and manufacturing, one expects that the implementation of design changes improves the product's reliability to achieve a design goal. Typically, a formal test analyze and fix (TAAF) program is implemented to discover design flaws and mitigate them. The gradual product improvement through the elimination of design deficiencies, which results in the increase of failure (inter)arrival times is known as *reliability growth* (RG).

Generally speaking, reliability growth can be applicable to any level of a design decomposition, ranging from a component to a complete system (Modarres et al. 1999). The (non-repairable) component level reliability growth can be readily established by comparing a chosen reliability metric for the consecutive design iterations or the product development milestone. For further reading on multiple comparisons of component level reliability data see Neslon (1995). Most of the existing reliability growth models, however, are associated with repairable systems (Crow 1974; Asher and Feingold 1984), so all mathematical notations in the remainder of the paper will be related to the repairable system treatment.

Today we find that RG models are often misunderstood, misapplied, and misrepresented. In many cases, the "internal customer" (e.g., product design engineer) is confused, and the credibility of the reliability practitioner is challenged or degraded. Critical business decisions may not be based on the most effective and appropriate reliability analysis available. Predictions are often constrained by data available from very small sample sizes (i.e., test hardware and test implementation cost constraints limit sample size far below typical "binomial based" approximation minimums for a given confidence level). This paper will attempt to communicate and characterize an understanding of the common and unique elements of both classical and Bayesian methods while assisting product engineering in the internalization, interpretation, and application of the appropriate growth models.

COMMON GROUND

Clearly, the fundamental principles of reliability growth method application and the intent to provide timely reliability feedback to product design and development are common to classical and Bayesian approaches. It is observed that both the classical and Bayesian RG models include depiction of a predetermined reliability target and are commonly used to compare designs and estimate reliability improvements for consecutive design iterations. Furthermore, both may include design verification (DV) tests, vehicle fleets, dynamometer test, bench test, system, subsystem, and component level test data as inputs to the corresponding model. They are consistent in producing a median value reliability growth curve along with the confidence/uncertainty intervals.

Mathematically, RG takes place, if the *rate of occurrence of failures* (ROCOF), denoted as $\lambda(t)$ in eq (1) is a monotonically decreasing function of time, t :

$$R(t_1, t_2) = \text{Exp} \left(- \int_{t_1}^{t_2} \lambda(t) dt \right) \quad (1)$$

CLASSICAL RELIABILITY GROWTH

Empirical studies conducted by Duane (1964) on a number of repairable systems have shown that the *cumulative mean time between failures* (CMTBF) plotted against cumulative time on test in a log-log space exhibit an almost linear relationship. The ROCOF, which is the inverse of the derivative of CMTBF according to the Duane model is

$$\lambda(t) = \alpha \beta t^{\beta-1} \quad (2)$$

where t is the total time on test, and α, β are the growth parameters.

Crow (1974) improved the Duane model by suggesting that equation (2) could be treated probabilistically as the ROCOF of a *non-homogeneous Poisson process* (NHPP). Such probabilistic interpretation of equation (2) is known as the Army Material Systems Analysis Activity (AMSAA) model and offers several advantages. First, the model parameters can be estimated through the maximum likelihood method and the confidence limits on these parameters can also be developed (Crow 1993). Second, the distribution of the number of failures $f\{N(t)\}$ can be obtained based on the Poisson distribution. It must be noted that the AMSSA model is not the only possible form for ROCOF of NHPP and some other models such as log-linear (Cox and Lewis 1966) can be used.

An important distinguishing feature of both Duane and AMSAA models is that the ROCOF is modeled as a (typically) decreasing function of time thus providing reliability growth according to equation (1). The models in question may be quite appropriate for monitoring progress of quality improvement of products in the field (Crow 1982, Asher 1984), but are much less effective to monitor reliability growth during the product development stage. Figure 1 shows a typical "growth chart" used to represent a classical RG.

Another RG model, which is sometimes also referred to as *classical* has to do with the so-called *design change credit* (DCC). Defense sector and automotive/private industry applications may include a complex subjective methodology to assess "design credits" (a.k.a., "fix effectiveness factor") to consistently make adjustments prior to engineering sign-off, production release, or "job 1" (Crow 1989, 1996). Under the DCC-based reliability growth model, the ROCOF in equation (1) is assumed to be constant thus implying that the failure occurrence is governed by the *homogeneous Poisson process* (HPP). While this assumption may be quite realistic in practice, it, nevertheless, provides no growth of reliability, unless the design change leads to the reduction of ROCOF:

$$\lambda'(t) = \varepsilon \lambda(t) = C,$$

where $\lambda(t)$ and $\lambda'(t)$ are the ROCOF of a product before and after the design change, respectively; $0 < \varepsilon < 1$ is a non-random quantity, which quantifies the DCC; and $C > 0$ is a constant.

Figure 2 shows a typical growth chart associated with the DCC-based reliability growth. While the DCC based RG model is subjective in nature and has no statistical grounds in estimating ε , it is accepted by management as a decision making tool in many industries.

Overall, one cannot deny that classical RG methods are useful product design management tools as demonstrated by the fact that they are highly accepted, well established and rooted in automotive/defense/aerospace industry sectors.

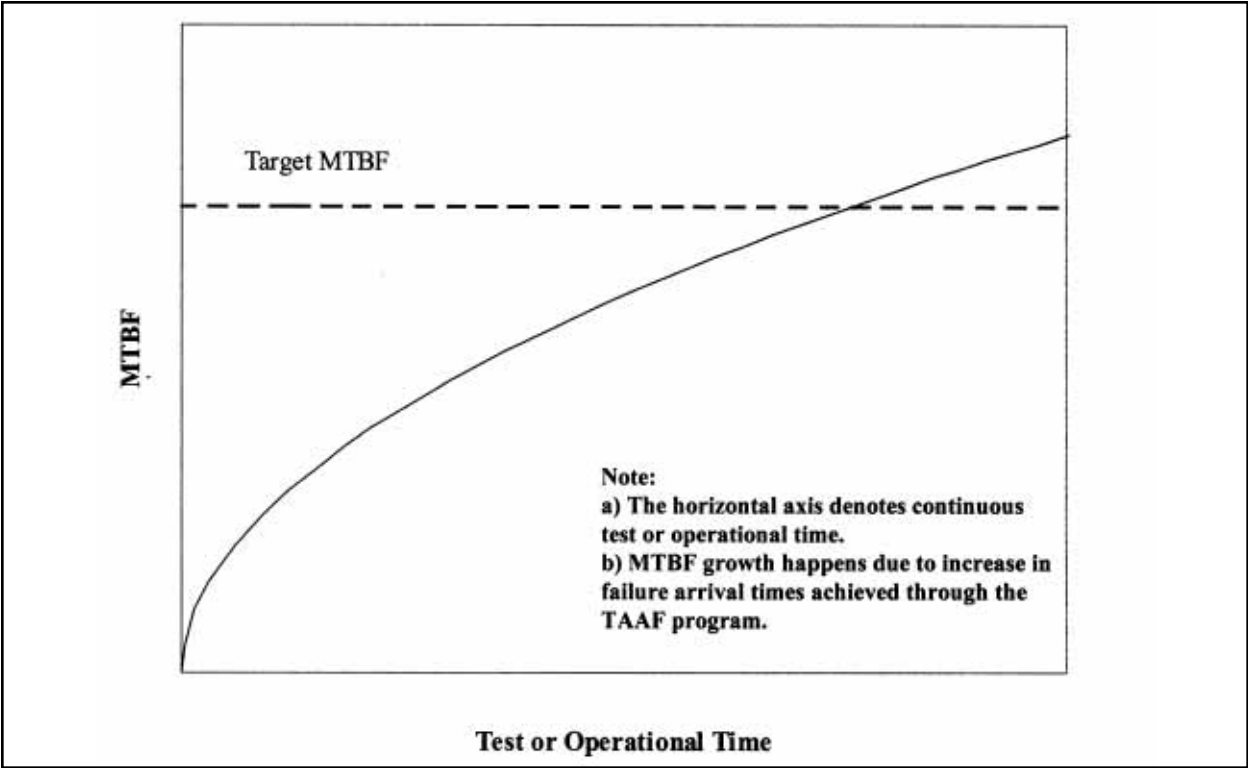


Figure 1. Classical reliability growth curve.

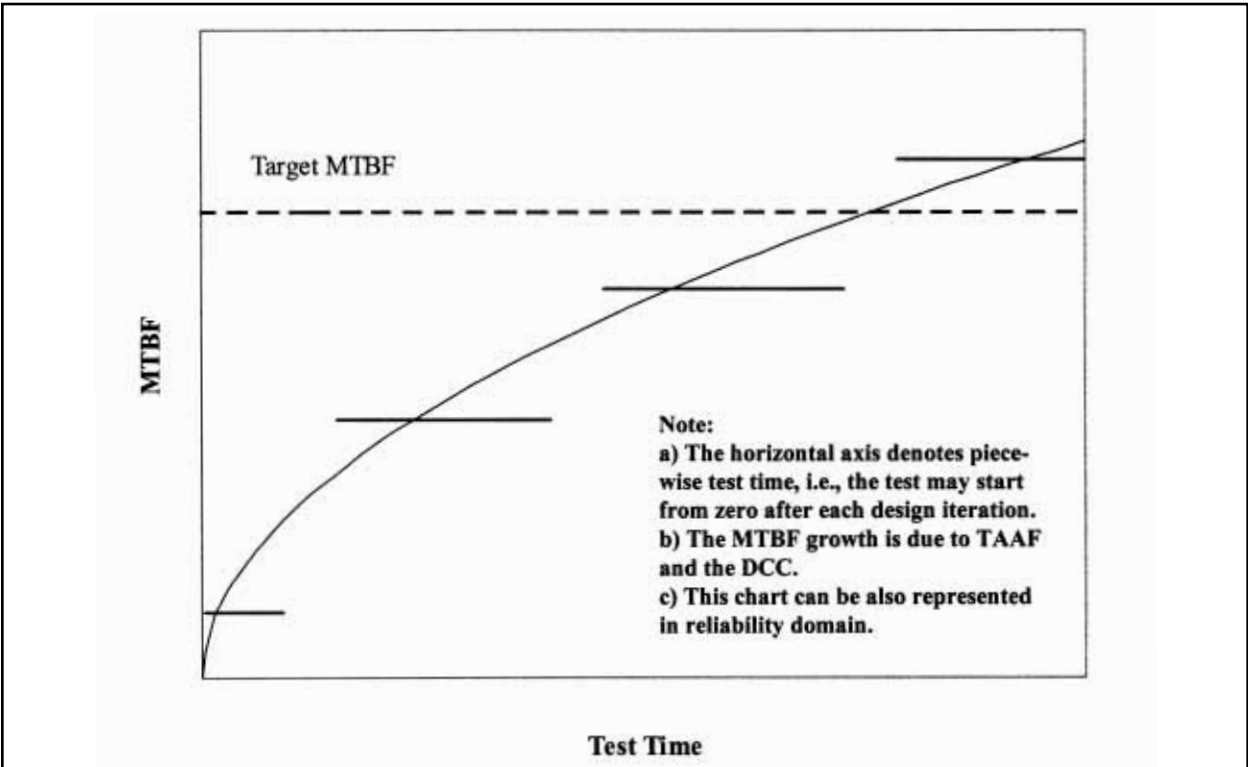


Figure 2. DCC-based reliability growth curve.

BAYESIAN RELIABILITY GROWTH

Bayesian-based reliability growth methods provide a wide range of reliability information relative to a specific design to support decision making in the product design process. The underlying theory guiding the process of estimating and updating product reliability is the Bayes' theorem (Martz and Waller 1982). As opposed to the classical approach whereby the growth parameters of (2) are treated as the unknown but fixed quantities, in the Bayesian framework these are interpreted as random variables, having a joint prior probability distribution, $\pi_0(\alpha, \beta)$. This distribution is then updated with available test evidence, E , in an iterative sequence to form the posterior distribution of the growth parameters, $\pi(\alpha, \beta | E)$:

$$\pi(\alpha, \beta | E) = \frac{\pi_0(\alpha, \beta) L(E | \alpha, \beta)}{\iint_{\alpha\beta} \pi_0(\alpha, \beta) L(E | \alpha, \beta) d\alpha d\beta},$$

where $L(E|\alpha, \beta)$ is the likelihood function, which represents the obtained evidence. The expected ROCOF is then computed as

$$\lambda(t) = \iint_{\alpha\beta} \alpha \beta t^{\beta-1} \pi(\alpha, \beta | E) d\alpha d\beta$$

to be finally used in equation (1) to estimate the achieved growth in reliability.

The main advantage of the Bayesian approach is that it can rely on multiple sources of evidence including: warranty data, customer research surveys, proving ground test data, etc. It also has the potential to systematically quantify and process "soft" evidence such as expert knowledge. Statistical evidence is ranked in terms of the degree of relevance to the design under study and categorized as either "partially relevant evidence" or "transformable evidence." Transformable evidence can be converted into a failure rate reduction factor or product life improvement factor, uncertainties of which are expressed as an interval, which includes optimistic, most likely and pessimistic values.

If desired by the practitioner, a "credibility interval" can be generated to display a lower and upper confidence bound in addition to the median estimate. Corporate memory including design verification information, test fleets,

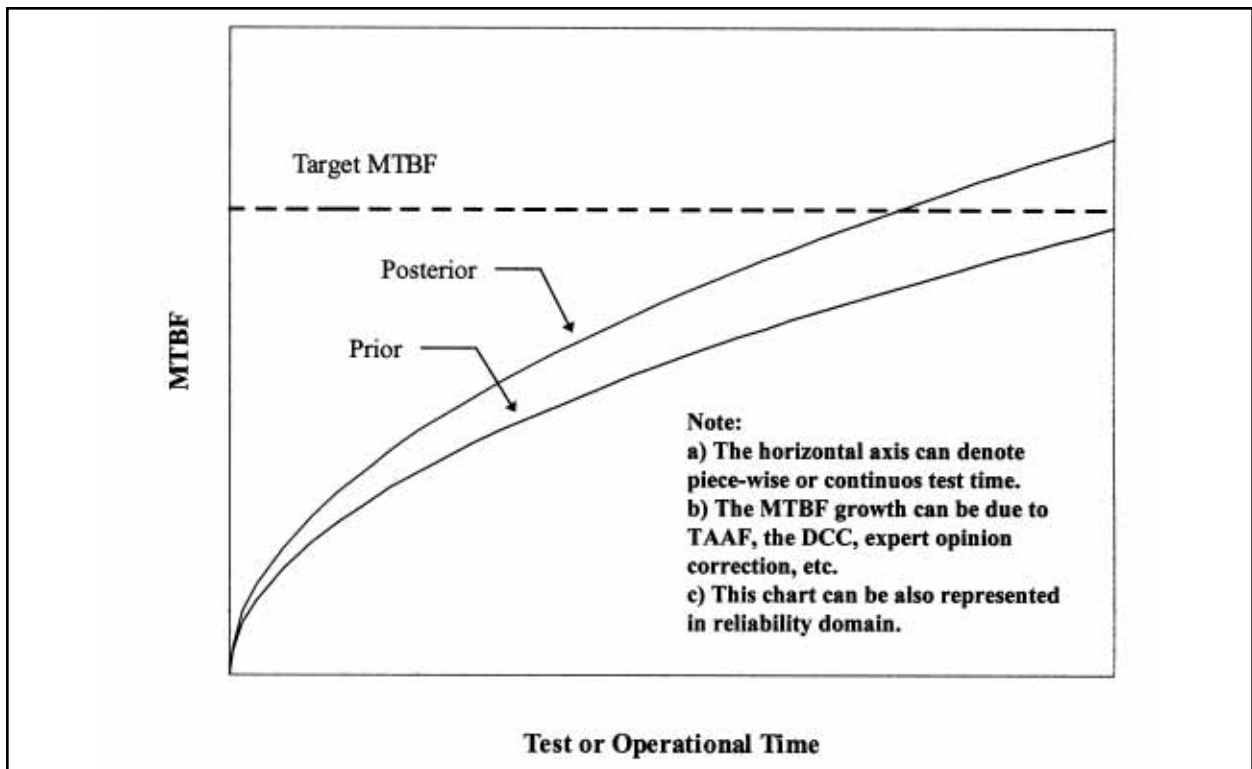


Figure 3. Bayesian reliability growth curve.

Table 1. Comparative analysis of the classical and Bayesian reliability growth methods.

	<i>Classical RG</i>	<i>DCC-based RG</i>	<i>Bayesian RG</i>
Advantages	well developed statistical models understood by management	simple to implement generally understood by management	meaningful use of the extensive prior data and test evidence (warranty, DV testing, expert judgement, etc.) can provide early (pre-test) assessment of reliability provide reduction in the uncertainty intervals resulting from a usually restricted test sample sizes
Disadvantages	rely on "hard" test/field data only under usually restricted sample sizes, provide wide confidence bounds	same as in classical RG weak statistical substantiation subjectivity in the design change credit assessment	if handled incorrectly, may bias the estimation results towards prior data based on a relatively involved mathematical theory, which makes it difficult for management's acceptance
Recommended use	in-field reliability monitoring evaluation of the repair/maintenance actions	reliability monitoring during the product development stage	same as in classical and DCC-based RG early, pre-test reliability assessment reliability assessment in the absence of the statistically valid test evidence

dynamometer test, laboratory/bench test, system, sub-system and component level test data can be easily integrated into the analysis approach. Figure 3 shows a typical growth chart representing a Bayesian reliability growth.

The caution that has to be exercised with the Bayesian approach is the degree of relevance of the prior data. If not assessed correctly, it may dominate in posterior estimation thus biasing the reliability status towards prior (historical) product reliability and away from the obtained test evidence on the current product.

CONCLUSIONS

Table 1 summarizes the reviewed RG models stressing the recommended areas of their application.

The reader may find it beneficial to establish generic reliability system and subsystem target assessment values (i.e., red, green, yellow) for use in health charts appropriate for Bayesian analysis. In effect, this would "normalize" the absolute results of the discussed methods and achieve equivalency in interpretation and communication of relative reliability estimation. Successful execution of classical and Bayesian methods will ensure realistic input is generated for effective decision making in product design and development.

REFERENCES

- Ascher, H., and H. Feingold. 1984. *Repairable Systems Reliability*. New York: Marcel Dekker, Inc.
Cox, D., and P. Lewis. 1966. *The Statistical Analysis of Series of Events*. London: Methuen.

- Crow, L. 1974. Reliability Analysis for Complex, Repairable Systems. In *Reliability and Biometry*. Ed. F. Proshan and R. Serfling. Philadelphia; SIAM, 379–410.
- Crow, L. 1989. Reliability Fix Effectiveness Factor Estimation. In *Proc. Annual Reliability and Maintainability Symposium*: 171–177.
- Crow, L. 1993. Confidence Intervals on Reliability of Repairable Systems. In *Proc. Annual Reliability and Maintainability Symposium*: 126–133.
- Crow, L. 1996. Reliability Growth Management, Models, and Standards. In *Proc. Annual Reliability and Maintainability Symposium*: 1–11.
- Duane, J. T. 1964. Learning Curve Approach to Reliability Monitoring. *IEEE Transactions on Aerospace* 2(2).
- Martz, H., and R. Waller. 1982. *Bayesian Reliability Analysis*. New York: John Wiley & Sons.
- Modarres, M., M. Kaminsky, and V. Krivtsov. 1999. *Reliability Engineering and Risk Analysis*. New York: Marcel Dekker, Inc.
- Nelson, W. 1990. *Accelerated Testing: Statistical Models, Test Plans, and Data Analyses*. New York: John Wiley & Sons.